

1 Human force control may trade-off 2 force error with central tendency 3 and recency biases.

4 **Hansol X. Ryu^{1,2*} and Manoj Srinivasan²**

*For correspondence:

hansol.ryu@ucalgary.ca (HXR);
srinivasan.88@osu.edu (MS)

5 ¹Biomedical Engineering, University of Calgary, Canada; ²Mechanical and Aerospace
6 Engineering, The Ohio State University, USA

7

8 **Abstract** Understanding how humans control force is useful to understand human movement
9 behaviors and sensorimotor control. However, it is not well understood how the human nervous
10 system handles different control criteria such as accuracy and energetic cost. We conducted
11 force tracking experiments where participants applied force isometrically while receiving visual
12 force feedback, tracking step changes in target forces. The experiments were designed to
13 disambiguate different plausible objective function components. We found that force tracking
14 error was largely explained by a trade-off between error-reducing tendency and force biases, but
15 we did not need to include an effort-saving tendency. Central tendency bias, which is a shift
16 towards the center of the task distribution, and recency bias, which is a shift towards recent
17 action, were necessary to explain many of our observations. Surprisingly, we did not observe
18 such biases when we removed force requirements for pointing to the target, suggesting that such
19 biases may be task-specific. This study provides insights into the broader field of motor control
20 and human perceptions where behavioral or perceptual biases are involved.

21

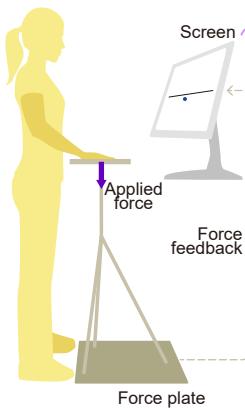
22 **Introduction**

23 We interact with the world by applying forces. For example, we push the ground to walk, press
24 pedals to ride a bicycle, and even apply muscle forces onto our own body parts to breathe, sing,
25 and reach out a hand. Understanding how people control forces would be beneficial in under-
26 standing how the nervous system performs sensorimotor control, insightful in designing robots
27 and biomechanical simulations. One way to model sensorimotor control is to view it as an opti-
28 mization problem, that there are objective functions that are optimized for and constraints that
29 are needed to be satisfied while executing the task (**Baron and Kleinman, 1969; Kleinman et al.,**
30 **1970**). There often are multiple objectives and constraints in many sensorimotor tasks and robotics
31 applications (**Dao et al., 2016; Jin et al., 2021**). Our goal was to observe human subjects' behaviors
32 when they perform simple force tracking tasks, and to investigate objective functions that could
33 explain the observed behaviors.

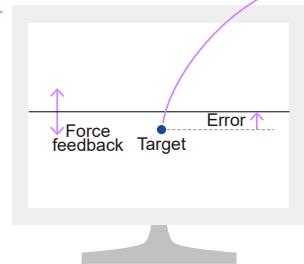
34 In biomechanics simulations and robotics, researchers often formulate an objective function
35 as a sum of error (or performance) term and effort (or energy cost, control energy) term. Such
36 formulations are used to model human motor control (**Emken et al., 2007; Izawa et al., 2008; Mi**
37 **et al., 2009**), in biomechanics simulations (**De Groot et al., 2016; Lee and Umberger, 2016**), and in
38 robotics (**Kalakrishnan et al., 2013; Miao et al., 2021**). In optimal control theory, cost function of a
39 linear quadratic regulator ("LQR", **Kalman et al. (1960)**) is often formulated as a sum of quadratic
40 function of state and control input.

Experimental design

(A) Experimental setup

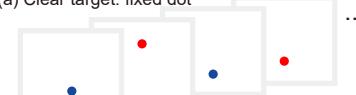


(B) Interface on screen

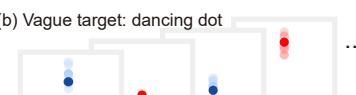


(C) Types of targets

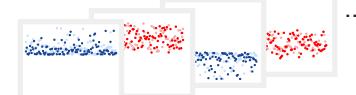
(a) Clear target: fixed dot



(b) Vague target: dancing dot



(c) Vague target: multiple dots



(D) Measurement example

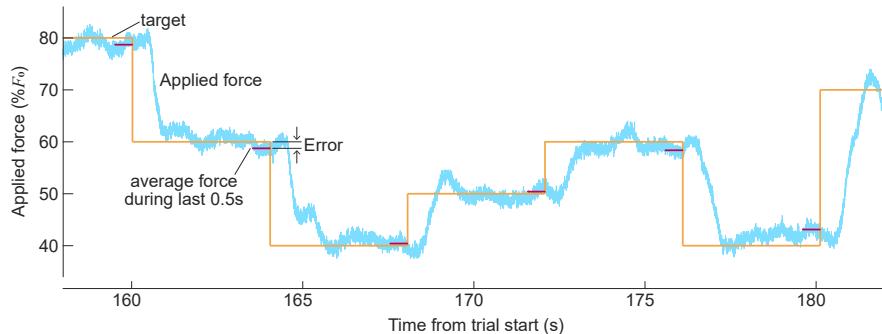


Figure 1. Experimental design. (A) Subjects applied force on the platform in front of them while looking into a screen. Vertical force they applied was measured through a ground-mounted force plate, and the force was relayed to a (B) screen that displayed applied force as a horizontal bar of a changing vertical location, as well as a target. (C) There were three types of targets. Illustrated are conceptual representations of each sub-trials. (D) Example of a trial, which was a series of sub-trials that changed target force in a random order. Applied force and target force at the time are shown together. Force tracking error was defined as a difference between the average applied force during the last 0.5 seconds of the sub-trial and the mode of the target.

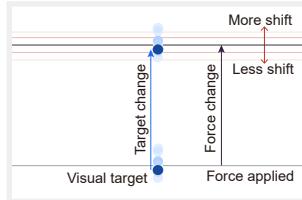
Having an error term in the objective function ensures that the goal is achieved to a certain degree. In addition to the error term, there are several reasons to include an effort term, including: 1) To handle redundancy problem: there are usually redundancy in the system, meaning there are multiple ways to achieve the goal, so additional criteria is needed (as discussed in *De Groot and Falisse (2021)*). 2) To better mimic biological systems: there are evidences suggesting that biological systems minimize energetic cost, so having the term allows simulation and robot to behave more like biological systems and thus provides more insights into understanding them (e.g., *Srinivasan (2011)*). 3) For practical reasons: Robots often need to be efficient because power or energy consumption are some of the major concerns (e.g., *Liu and Sun (2013)*; *Pellegrinelli et al. (2015)*).

However, those functions and their relative weightings are usually arbitrarily designed and tuned until they produce acceptable outcomes. Some researchers refer to biological measurements to formulate some components of the objective function (e.g., *Kording and Wolpert (2004)* studied error, *Berret et al. (2011)* studied cost), but such investigations have not been considered in the context of human force tracking.

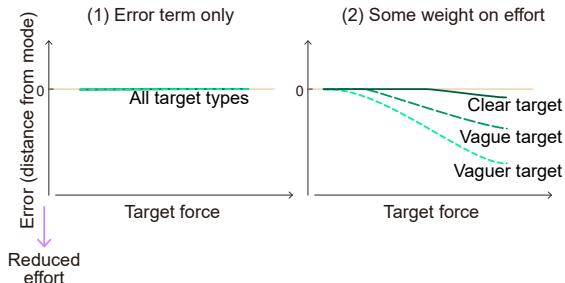
We designed a force tracking experiment that allows us to compare various objective function models to experimental observations. Participants were asked to apply forces isometrically onto the platform in front of them through their hands, while receiving the force feedback as the height

Expected results

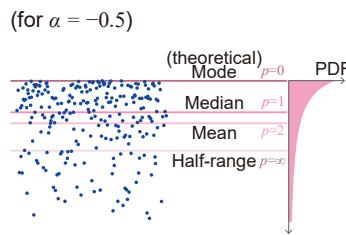
(A) Target change vs. force change



(B) Conceptual illustration: effect of effort term on overall error-force relationship



(C) Various centers of a distribution



(D) Conceptual illustration: effect of effort term on error-distribution parameter relationship (for 1 < p < 2)

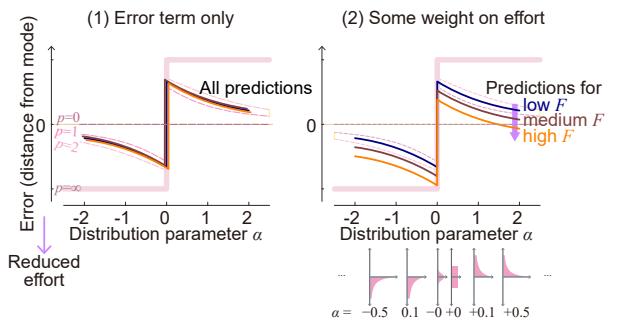


Figure 2. Illustration of expected results for different target types and distributions. (A) Relative force shifts with respect to the target shift. Subjects may shift by the same degree as the visual target position changes, but they may shift less than the ascending target change if effort becomes a critical factor. (B) Therefore, if minimizing error is the only factor that determines people's behaviors, (1) we expect average zero error for all force ranges and target types. However, if effort is a consideration too, and if vagueness of the target as an extra allowance to facilitate effort-saving behavior, (2) we expect increasing error towards the effort-saving direction as target vagueness increases. (C) Using targets of different distributions could provide us extra information about how people perceived the task. As there are many definitions of centers in a skewed distribution, including the mode, median, mean, and half-range, we could study where people were aiming for, and whether this changes with increasing force demands. (D) Combined with effort considerations, if effort does not affect the behavior and people are consistent with what "center" they aim for, (1) we expect to see error trend across various distribution parameters that is best described by which center they aimed for, and does not change with increased force requirements. If effort is combined with this perception, (2) we expect that people's response will be similar to that of (1) when force requirement is low, but will shift towards effort-saving way as force requirement increases.

58 of the bar on the screen (figure 1A, B). The force targets (figure 1C) they tracked were either clearly
 59 defined as a fixed single dot, or vaguely defined in two ways, as a single dot that had noise in its
 60 location or as a cloud of multiple dots. The targets were shown for 4 seconds (referred to as "sub-
 61 trial"), and the next target of the same kind appeared at a random location. We analyzed force
 62 error near the end of each sub-trials (figure 1D) to quantify the force tracking error.

63 We hypothesized that the vagueness of the target would result in more shifts in subjects' be-
 64 haviors towards the effort-saving direction when force requirement increases. If effort is not a
 65 criterion, when subjects see the same distribution of dots simply shifted up to require more force
 66 (figure 2A), they would increase their force by the same amount. In that case, their behaviors would
 67 be largely explained by an error-based objective function alone (figure 2B-1). However, as there is
 68 some allowance due to the vagueness of the target, they may save effort by increasing the force
 69 by a slightly smaller amount than the target shift. In that case, we could model the behavior as a
 70 trade-off between error-reducing and effort-saving criteria in the objective function (figure 2B-2).

71 In addition, we used skewed distributions when we defined vague targets to study error term

72 of the objective function (figure 2C). Skewed distributions have distinct mode, median, mean, and
73 half of the range, which are the locations where error raised to power of 0, 1, 2, and infinity are
74 minimized. For example, if a subject minimized error raised to about 1.5, on average they would
75 place the force feedback bar between the median and mean of the distributions (figure 2D-1). If an
76 effort-saving tendency is added to the same error-reducing criterion in the objective function, we
77 expect that people's behaviors would not change much in low force range, but it will shift towards a
78 direction that saves effort when the force requirement increases (figure 2D-2). We measured how
79 participants changed their force according to the force requirement and given distributions, and
80 applied inverse-optimization analysis to look for an objective function that best describes observed
81 human behaviors.

82 We observed shifts towards less forces in medium-high force range. However, there also were
83 unexpected findings that looked like people systematically "wasted" effort in low to medium force
84 range, meaning, they exerted more force than needed for no obvious benefits in reducing errors.
85 We used "central tendency bias" and "recency bias" to explain the results, and conducted additional
86 follow-up experiments on a subset of the participants to test these biases. We use the term "central
87 tendency bias" to describe the shift towards the center of the tasks. That is, if their current force is
88 on a lower side, people tend to make errors towards the center, and end up producing more force
89 than needed. In a similar manner, we use the term "recency bias" to refer to the bias towards the
90 recent past action, meaning people tend to make an incomplete shift to the next target. Surpris-
91 ingly, an effort term that increases with force, does not seem to have a significant effect on the
92 model prediction. Further, we found that eliminating the force requirement as much as possible
93 eliminated the biases, showing no significant bias in visual perception of the tasks.

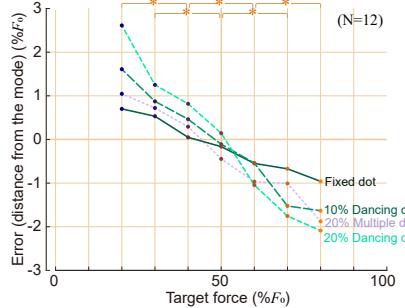
94 **Results**

95 **How humans track step changes in forces**

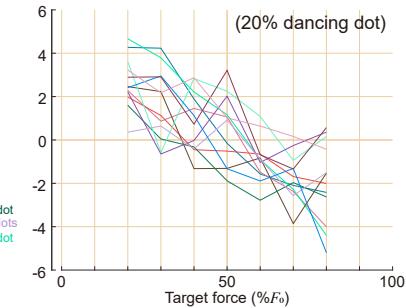
96 Force tracking errors have task-dependent positive and negative biases

Results1: force tracking error as a function of force and target type

(A) Results: overall error-target force



(B) Results: Individual subject



(C) Results: error - distribution parameter relationship for different target forces

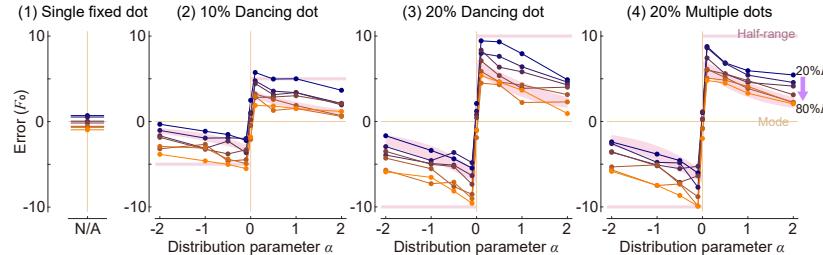


Figure 3. Median values of tracking error for different target force, target types and distribution parameters.

(A) Error, defined as a distance between average force people applied during the last 0.5 seconds of sub-trials and a mode of the target, for each target types. Asterisks indicate statistically significant difference ($p<0.05$) for pairs of the force conditions, for every target types compared separately. (B) Error-force pattern of individual subjects. (C) Median response for each target force and distribution parameter. Pink shaded region represents the range between theoretical mean and median of each distributions, and pink horizontal line represents the half of the distribution range.

97 Subjects typically took about 0.4 s seconds to initiate the force change after target location changed,
98 and held the force until the next target was shown (figure 1D; average responses can be found in
99 figure 9 and figure 10). We report distance from the mode of the distribution as the error in this
100 paper unless otherwise noted.

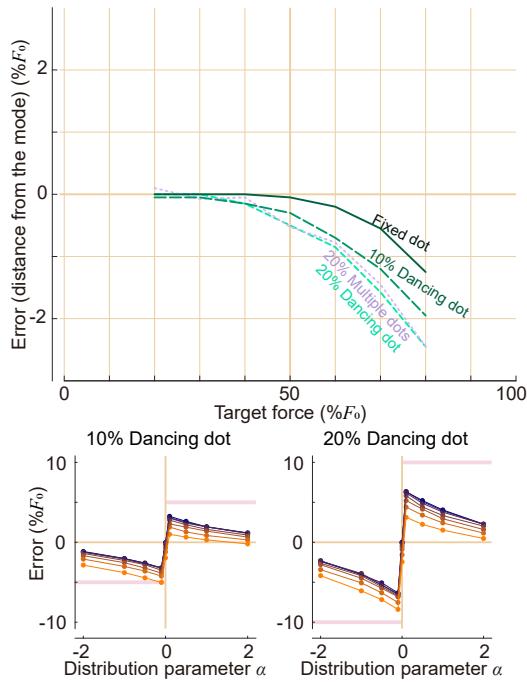
101 We measured the force tracking error for different target types and force requirements. As we
102 expected, people exerted less force relative to the target when the target force increased, and there
103 was a bigger change when the target was vaguer (figure 3A). However, what was unexpected for us
104 was that people tended to exert more force than needed for low-medium range targets. These two
105 observations, 1) negative correlation between error and target force and 2) positive error on the
106 lower target force, were consistent across subjects (figure 3B). When targets were grouped based
107 on the distribution parameters, force errors were in general in between median and half-range
108 of the distributions. Each of the sub-grouped forces errors showed a negative correlation with a
109 target force (figure 3C).

110 Force tracking error results are not explained by error and effort minimization

111 We used two types of objective function models to describe the observations. The objective func-
112 tion we originally formulated was in the form of $\sum \text{error}^p / N + cF^\gamma$, where error is absolute distance
113 between the target and applied force, F is applied force, N is number of target dots. Hyperpa-
114 rameters p and γ are shape parameters for each functions, and c is a constant that determines

Competing models of force tracking error

(A) Error & effort model



(B) Error & bias model

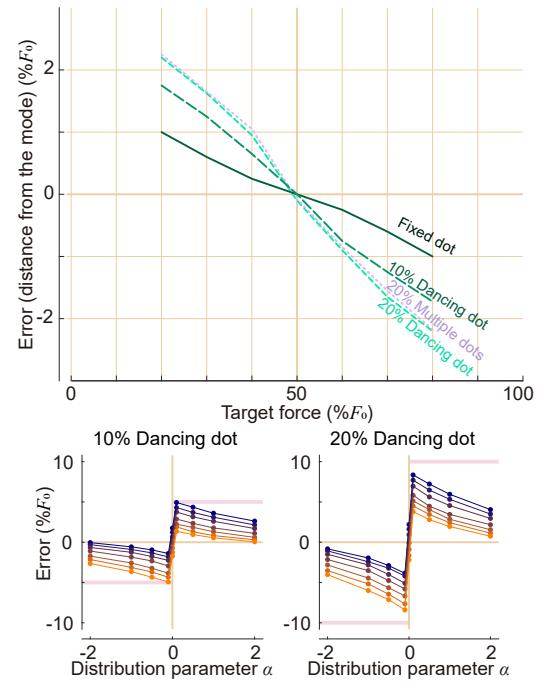


Figure 4. Modeling force tracking errors using (A) error-and-effort model and (B) error-and-bias model.

In error-force plots (top row), different target types are indicated with different line types. In distribution parameter-error plots (bottom row), lighter and more orange colors indicate lower target force, darker and more blue colors indicate higher target force. Error shown on y axis in both types of plots are both error with respect to the mode of the target distribution. Objective function models to generate the model prediction here are $\sum \text{error}^{1.6}/N + 0.06F^5$ for (A) and $\sum \text{error}^{1.6}/N + 0.15 d_{\text{ctr-bias}}^{1.8}$ for (B).

115 the relative weighting between two terms. Since human data is noisy, we aimed to keep the for-
 116 mulation as simple as possible. For the simplicity, we used $\sum(\text{error}^p)/N$ to represent performance
 117 criteria for targets of different vagueness's. The formulation has shallower minima for bigger N if
 118 $p < 2$, thus competing terms in the objective function would make the minima shift more for bigger
 119 N . The effort term in this form has inherent limitations that it cannot explain positive force errors
 120 (figure 4A), and thus inverse optimization did not have a local minima within our search range. We
 121 illustrated a solution that was found on one end of the search range, but even bigger exponent
 122 on effort term would further minimize the objective function. Bigger exponent would make low to
 123 middle range error closer to zero, thus reduces model error. However, this formulation does not
 124 predict positive error in any case.

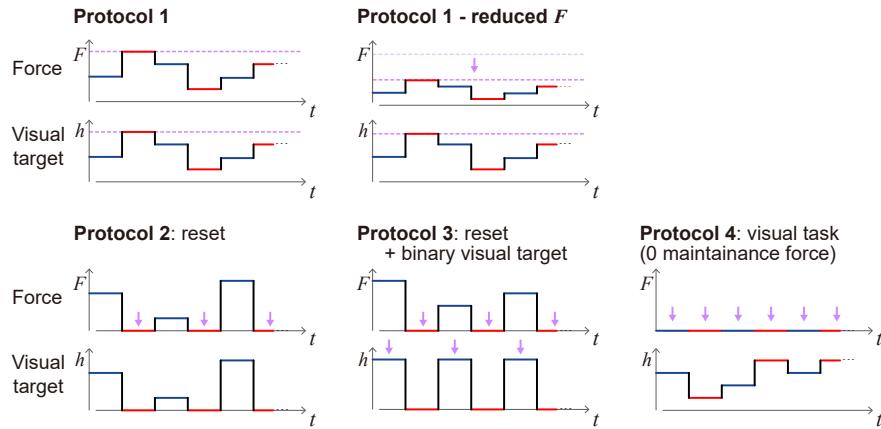
125 We investigated which variable predicts force tracking error.
 126 Force tracking error might be fully determined by the target force or visual representation of the
 127 target force. However, there might be a bias that is determined by the temporal relationships
 128 between targets. We formulated variants of the objective function models to represent these cases.
 129 One of our objective function models was error-and-bias model (figure 4B), which had a form of
 130 $\sum \text{error}^p/N + bd_{\text{bias}}^\beta$, where d_{bias} represents the distance to some point when there is a tendency to
 131 shift towards that point, which will be further explained later. error is absolute distance between
 132 the target and applied force. Hyperparameters p and β are shape parameters for each functions,
 133 and b is a constant that determines the relative weighting between two terms. figure 4B shows the
 134 error-and-bias model that has a shift towards the average target force.

135 **Additional experiments to disambiguate different terms**

136 Some of the variables inherently co-varied in the original protocol (referred to as "Protocol 1"), so
137 we designed additional experiments to distinguish some of them and test models accordingly.

Additional experiments to test various models

(A) Additional test protocols



(B) Model predictions based on measurements from protocol 1

(1) Error & effort model

(2) Error & central bias model

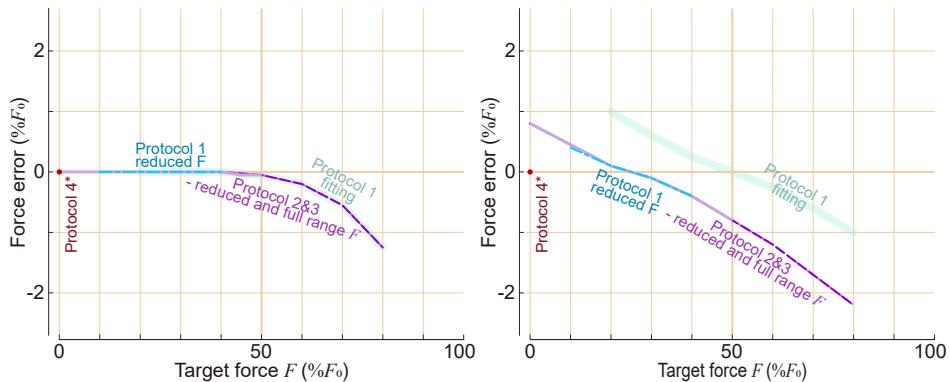
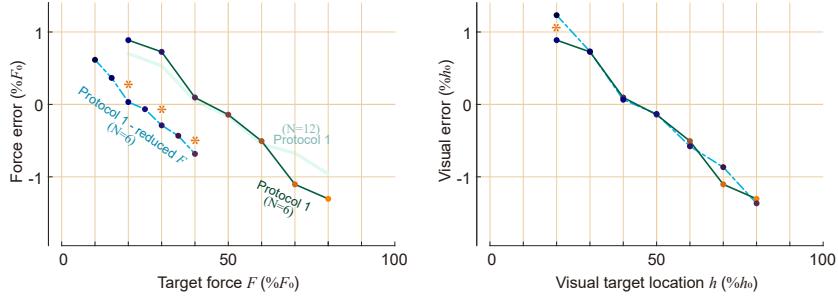


Figure 5. Additional protocols to test different model predictions. (A) Test protocols in terms of force requirement and its visual representation on the screen. Purple arrows emphasize the main differences between protocols. Protocol 1-reduced F tested for a lower half range of the force, while spanning to a full visual target range. Protocol 2 and 3 had zero re-set period between each non-zero target forces. Protocol 3 had binary targets, where the target was alternating between only two locations each indicating zero and non-zero target. This was done by changing the conversion ratio between force and visual target location each time. Protocol 4 was done using a computer mouse, and did not require force to maintain the pointing location. (B) Example of model predictions for these additional protocols. (1) Error-and-effort model, as an example of a model that target force determines the error: results from all protocols will match with each other when the target force is the same. (2) Error-and-bias model, as an example of a model that depends on a target force and the experimental contexts. Error-force trend changes for each protocols, and we had specific qualitative expectations for each protocols.

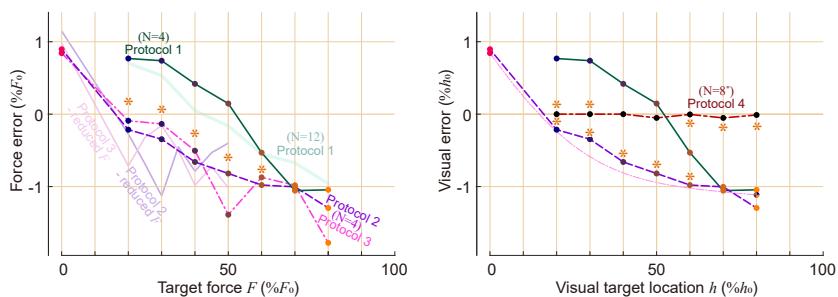
138 Design of the additional protocol and what we found from each of them are listed below. A
139 subset of subjects who participated in original experiment using Protocol 1 participated in these
140 additional tests, so we compared their responses to those of Protocol 1. We used a single fixed
141 dot as a target in all additional tests to reduce the variability due to the target vagueness.

Results2: additional protocols

(A) Experimental results: protocol 1 reduced range



(B) Experimental results: protocol 2, 3, 4



(C) Experimental results vs. model

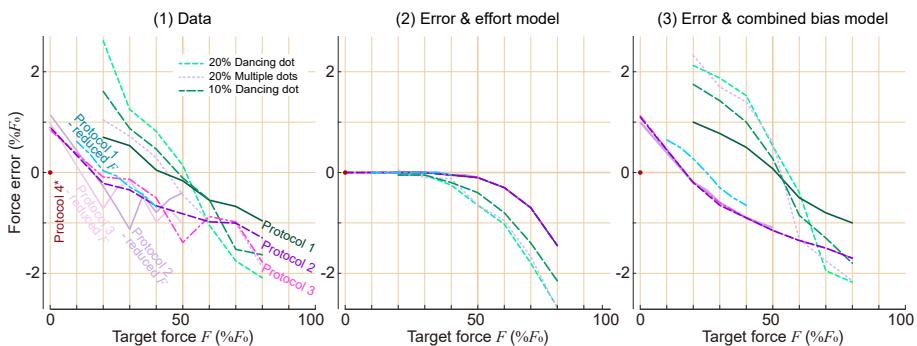


Figure 6. Experimental results of additional protocols and its comparison to model predictions.

Different lines indicate results of each protocols. As subsets of Protocol 1 subjects participated in additional tests, we represented the results of matching subjects in Protocol 1 as a thick green line here, while showing the results of all subjects as a dimmer line. Matching subjects result of Protocol 4 was omitted for simplicity, because the error was consistently very close to zero for everyone. (A) Protocol 1-reduced F trial had a distinct error trend compared to Protocol 1 in force domain, while its visual domain had a good match. Asterisks indicates statistically significant different ($p < 0.05$) difference between Protocol1-reduced F and Protocol 1 for matching conditions. (B) Protocol 2, 3, 4 results in force and visual domain. Asterisks indicates statistically significant difference between protocols and Protocol 1. (C) Comparison of (1) experimental results and (2, 3) models. Error-and-effort model and error-and-effort models were poor and good predictor of data.

142 **Protocol 1 – reduced F** shows that target force alone does not determine error.
143 In Protocol 1 – reduced F , we tested lower half range of the force while keeping the range of the
144 visual target fixed (figure 5A). Among our 12 subjects who participated in Protocol 1 experiment, 6
145 of them participated in this additional protocol experiment on the same day.
146 **Results:** Testing result shows that force-error relationship was significantly different between Pro-
147 tocol 1 and Protocol 1 – reduced F , but visual target location – error relationship in general matched
148 quite well between two protocols (figure 6A).

149 **Protocol 2 and 3** shows that testing order affects the error, and that visual representa-
150 tion of the target force is not a good descriptor of error.
151 We tested the same target force range as Protocol 1, but added zero force reset periods between
152 non-zero force targets (figure 5A) to test the order effect. The difference between Protocol 2 and
153 Protocol 3 was the visual representation of the target. In Protocol 2, visual representation was the
154 same as the Protocol 1, i.e., target location was moved proportional to the required force. In Pro-
155 tocol 3, targets appeared between only two locations, lower one indicating 0 force and higher one
156 indicating various non-zero target forces. In this case, the ratio between force and target repres-
157 entation changed each time new target was presented. On a different day from Protocol 1 testing, 4
158 participants came back to participate in Protocol 2 and Protocol 3.
159 **Results:** Testing results suggest that visual representation of the target is also not a good predictor
160 of the force tracking error. Despite different visual target representations, there wasn't significant
161 difference between Protocol 2 and Protocol 3 in force-error relationship on all target force condi-
162 tions (figure 6B). Those force-error relationships were significantly different from that of Protocol
163 1 in the most of the target force range. The observations were similar in visual target location –
164 error relationships. Error trends of Protocol 2 and Protocol 3 were similar to each other, but were
165 different from that of Protocol 1. We also tested Protocol 2 - reduced F and Protocol 3- reduced F
166 conditions, where we tested smaller range of the forces, while not changing the mapping between
167 force and its visual representations. There were no notable differences between Protocol 2 and 3
168 and their reduced F versions.

169 **Protocol 4** shows that visual bias does not explain the force bias.
170 We asked subjects to track the target using a computer mouse, so that we could investigate the
171 errors coming from visual aspects of the task. They did not need to keep exerting force to keep
172 the mouse position, which is different from all other protocols where they needed to keep apply-
173 ing the force until new target was shown. Eight out of twelve subjects performed this Protocol 4
174 experiment on their personal computers.
175 **Results:** Participants performed the tracking task almost perfectly. Visual tracking errors were
176 overall quite small, and the error relationship was again significantly different from that of Proto-
177 col 1 for most of the ranges (figure 6B).

178 Model with force error and biases captured the experimental results qualitatively
179 We reject hypotheses that force or visual representation of the force predict force tracking error,
180 as force tracking errors were different for some protocols even for the same force or same visual
181 target location. Any objective function model that mainly rely on mere force (e.g., figure 6C2) or
182 visual representation to describe error trend would not be suitable to describe the observation.
183 Error-and-bias model was better supported by the data. Error-and-central tendency bias model
184 that fits Protocol 1 predicts the results of additional protocols fairly well (figure 6C3). The objective
185 functions that minimized mean RMS error between mean data and mean model prediction on each
186 force condition of different protocols were as follows (figure 7):

187 • Error-and-central tendency bias model:

$$\begin{aligned} F^* &= \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.7}}{N} + 0.09d_{\text{ctr-bias}}^{1.55} \\ &= \arg \min_F \sum_{j=1}^N \frac{|F_{\text{target},j} - F|^{1.7}}{N} + 0.09(\bar{F}_{\text{target}} - F)^{1.55}, \end{aligned}$$

188 • Error-and-recency bias model:

$$\begin{aligned} F^* &= \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.5}}{N} + 0.15d_{\text{rec-bias}}^{1.3} \\ &= \arg \min_F \sum_{j=1}^N \frac{|F_{\text{target},j} - F|^{1.5}}{N} + 0.15(F - F_0)^{1.3}, \end{aligned}$$

189 • Error-and-combined bias model:

$$F^* = \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.6}}{N} + 0.05d_{\text{ctr-bias}}^{1.55} + 0.07d_{\text{rec-bias}}^{1.3},$$

190 where F^* is the prediction of the applied force based on the given objective function and F is the
191 applied force that is to be optimized. $F_{\text{target},j}$ is the force indicated by j -th dot in the sub-trial, when
192 N is the number of dots presented. Bias-related terms $d_{\text{ctr-bias}}$ is a distance to the "center" of the
193 targets associated with central tendency bias, $d_{\text{rec-bias}}$ is a distance to the recent action associated
194 with recency bias. They are determined by \bar{F}_{target} , which is the average (or expected) target force
195 of the whole trial consisted of many sub-trials, and F_0 , which is the force at the beginning of the
196 sub-trial.

197 Details of the inverse optimization analysis is described in the Methods section.

Variants of error-and-bias models describing the force tracking error

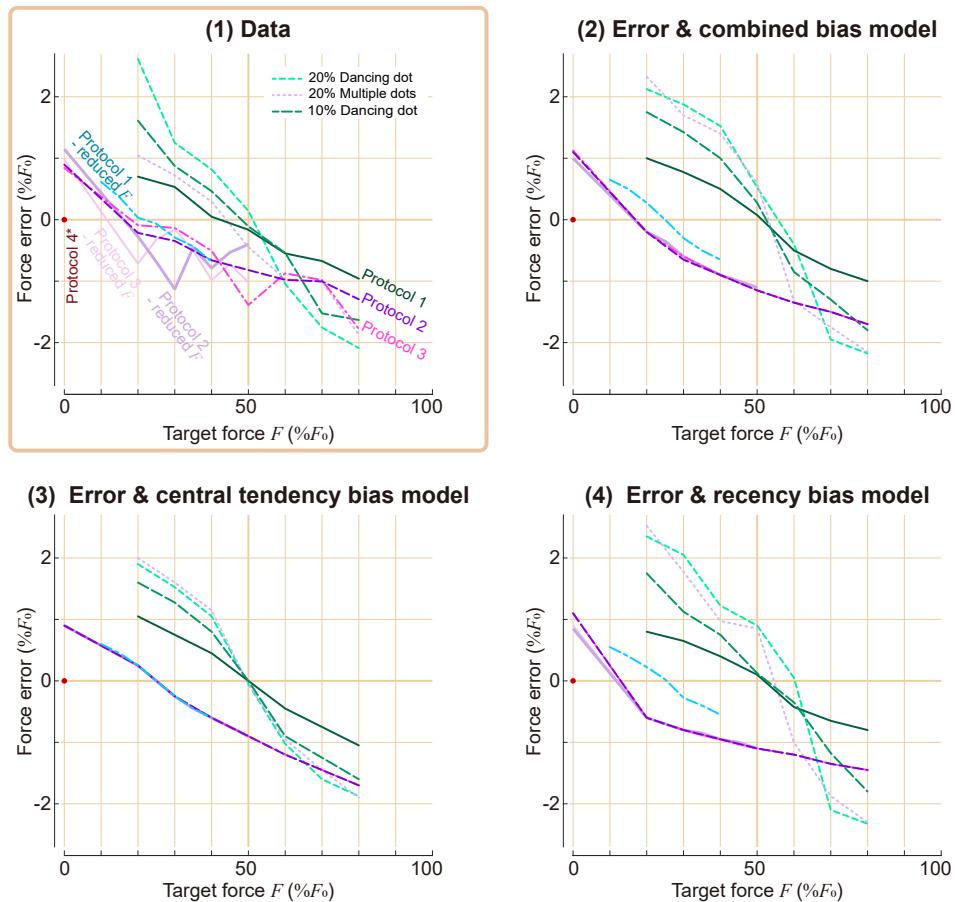


Figure 7. Variants of error-and-bias models describing the force tracking error from various protocols.

(1) Experimental data from various protocols. Different lines indicate median results of each protocols. (2)-(4) Models that best fit the overall experimental data, whose objective functions are weighted sum of error term and (2) both central tendency bias and recency bias terms, (3) central tendency bias term, and (4) recency bias term.

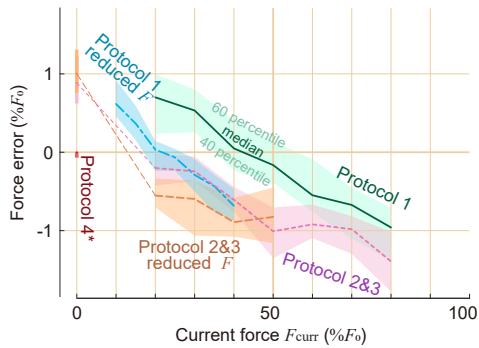
198 **Changes in target forces describe error trends from all protocols.**

199 We further investigated how force tracking error is explained by a tendency to shift towards some
 200 point. As we pointed out earlier, the error-target force relationship (figure 8A) and the visual error-
 201 visual target location relationship had distinct patterns for each of the protocols. However, ex-
 202 pected change in target force (figure 8B) and change in target force from the recent past target
 203 (figure 8C) better describe error across different protocols. These support the idea of people's
 204 tendency to shift towards either the recent past action or the center of the task. Error-and-effort
 205 model could not explain the positive error, but error-and-bias models could explain it: central ten-
 206 dency bias predicts positive error on a lower-range target forces as the center of the task is higher
 207 than the current force, and recency bias predicts positive error on lower range target forces as
 208 their recent past targets were on average higher than the current target.

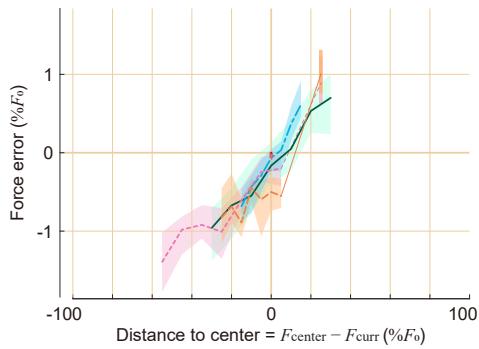
209 We defined recency bias as a function of $d_{rec-bias}$, which is the distance to the recent past action,
 210 and central tendency bias as a function of $d_{ctr-bias}$, which is the distance to the "center" of the targets
 211 or the average of the targets. From their definitions, the average of the $d_{rec-bias}$ is close to $d_{ctr-bias}$
 212 except for the small discrepancies due to the order effect in Protocol 2 and 3. Therefore, models
 213 using recency bias, central tendency bias, or combination of recency bias and central tendency bias,
 214 all yield similar overall error trends (figure 7C). We analyzed sub-grouped data to see their separate
 215 effects, and found that they both seem to exist, which is presented in the following section.

Error with respect to force, distance to force center, force distance to recent target

(A) Error - current force



(B) Error - distance to center (central bias)



(C) Error - distance to recent target (recency bias)

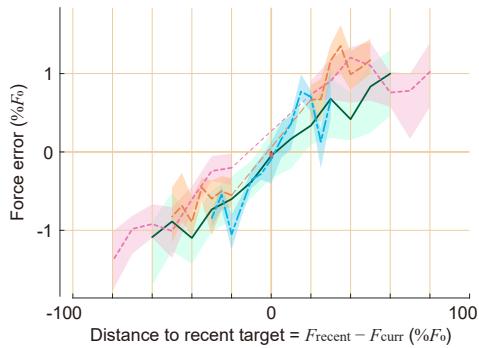


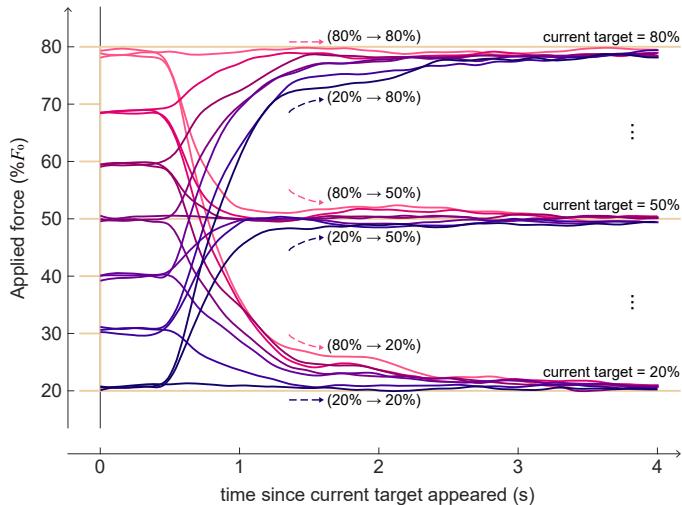
Figure 8. Error trends with force, force transition from recent target, and expected future force transition. Force tracking error is plotted against (A) the current target force, (B) distance to the target center from the current target force, which is the expected transition, (C) distance to the recent past target force from the current target force, which is the transition from the past. Shaded areas represent 40 to 60 percentiles of the data at each condition, and the thick lines represent the median of the data. We showed visual error of Protocol 4 on the y axis as a reference, because force error is undefined. Y axis is for force error for all other conditions.

216 **Comparing central tendency bias and recency bias**

217 Changes in target force affects error when current target force is fixed.

Recency bias: effect of target force change ΔF

(A) Average force transitions, grouped by current target force F_{curr}



(B) Data and bias models describing error- ΔF relationship

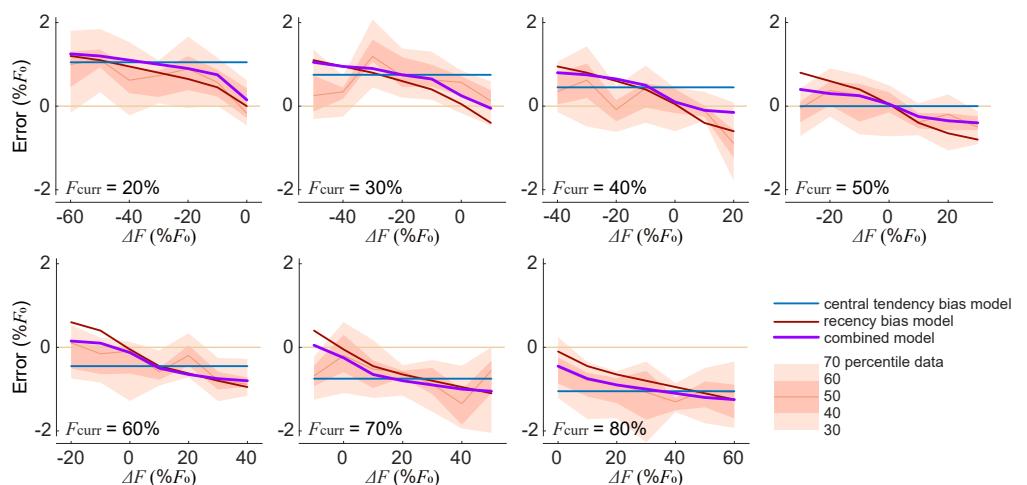


Figure 9. Protocol 1 sub-trials grouped into same target forces. (A) Averaged force time series grouped based on its recent past target and current target forces. Brighter and more red colors are for higher previous target force, and darker and more blue colors for lower previous target force. (B) Force error as a function of target force change, among sub-trials that had same current target force. Each sub-panels are for different current target forces. Shaded area represents 30, 40, 50 (=median), 60, 70 percentile data range. Blue, red, purple lines are predictions using error-and-bias model with only central tendency bias, only recency bias, or central bias and recency bias combined.

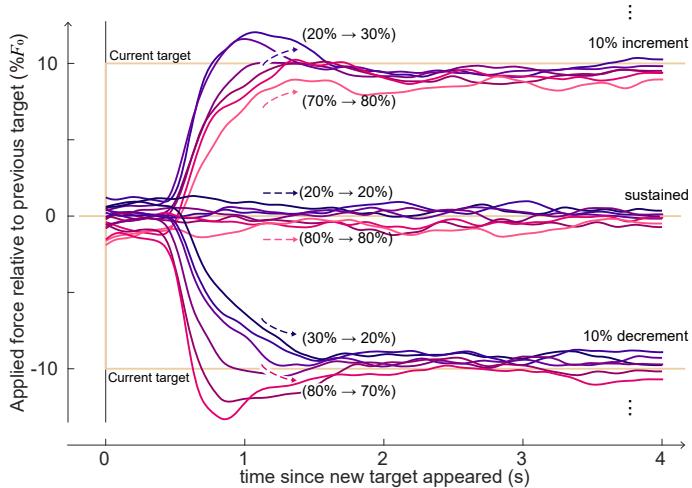
218 We grouped data of Protocol 1 – fixed dot target condition to investigate the recency bias effect. If
 219 recency bias does not have a notable effect, sub-trials that have the same target force will generally
 220 have the same errors, regardless of its recent past target. However, we do see a dependency on
 221 where the force started from, even if the target force is the same (figure 9A). Errors had negative
 222 correlations with force change, even when they are compared against sub-trials that had the same
 223 target force.

224 We considered three types of bias models in the objective function: central tendency bias, re-
 225 cency bias, and a combined model that has both biases. While all three of them do similarly well at
 226 capturing the overall error trend with force, central tendency bias model does not capture the neg-
 227 ative correlation between force change and error when the target force is fixed (figure 9B). Recency
 228 bias and combined model are more suitable to describe this observation.

229 Current target force affects the error among sub-trials that underwent same change in
 230 target force.

Central tendency bias: effect of current force F_{curr}

(A) Average force transitions, groupd by force change ΔF



(B) Data and bias models describing error- F_{curr}

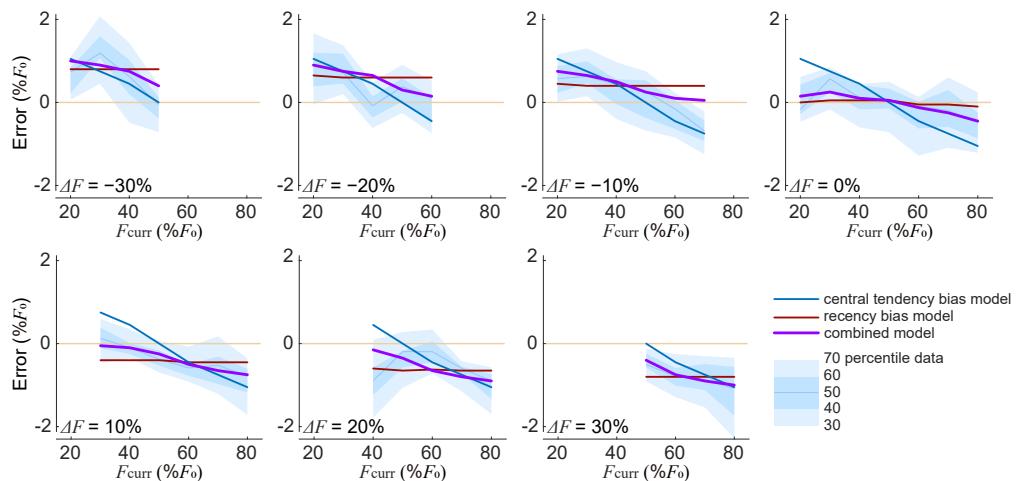


Figure 10. Protocol 1 sub-trials grouped into same changes in target forces. (A) Averaged force change time series, where y axis is a force difference with respect to the recent past target. Shown here are sub-trials that had $10\%F_0$ increment, same, or $10\%F_0$ decrement from the recent past target. Brighter and more red colors are for higher current target force, and darker and more blue colors for lower current target force. (B) Force error as a function of current force, among sub-trials that had same change in the target force. Each sub-panels are for different changes in the target force. Shaded area represents 30, 40, 50 (=median), 60, 70 percentile data range. Blue, red, purple lines are predictions using error-and-bias model with only central tendency bias, only recency bias, or central bias and recency bias combined.

231 We now grouped data of Protocol 1 – fixed dot target condition to investigate the central tendency
232 bias effect. Here, we grouped sub-trials into the same target force changes, because analysis on
233 the sub-trials that had same target forces showed that change in target force affects the force error.
234 If recency bias alone could explain the observations and central tendency bias is has a negligible
235 effect, groups of data that underwent the same target force change would have a similar error
236 regardless of the current force requirement. However, we do see a dependency on the current
237 target force level, when we compare sub-trials against other sub-trials that had the same target
238 force change (figure 10A). Errors had negative correlations with target force in each of the groups
239 that had same change in target forces. We explained this observation in terms of central tendency
240 bias, that people make positive error when the center of the task is higher than the current target
241 force, and make negative error when the center is lower than the current target force.

242 We again considered three types of bias models in the objective function: central tendency bias,
243 recency bias, and a combined model that has both biases. Since change of target force is a variable
244 that inherently co-varies with target force on average, all three models capture the overall error
245 trend. However, recency bias model does not capture the negative correlation between target
246 force and error when the target force change is the same (figure 10B). Central tendency bias and
247 combined model are more suitable to describe this observation.

248 **Purely visual tasks did not have the biases**

249 We have showed that force or visual target locations alone were not good descriptors of error, but
250 force bias could be a good descriptor. Then, if error could be explained by force bias, there is a
251 chance that it could as well be explained by a visual bias. We therefore investigated whether visual
252 bias could be an alternative explanation of the error trend we observed. Since visual location of
253 the target was proportional to the force in Protocol 1, visual bias could explain the error trend
254 (figure 11 left) as much as the force bias does. The most striking counterexample is from Protocol
255 4, where subjects used a computer mouse to track the targets and did not need to exert force to
256 maintain the pointing location. Although recency bias is commonly studied in visual tasks, we do
257 not observe clear visual biases from Protocol 4 (figure 11 right). Subjects made almost no error
258 for single fixed target when they use a computer mouse to track the target. When they tracked a
259 vague target which is a set of multiple dots, overall variability increased, but there was no significant
260 trend in visual error, and its mean is not significantly different from 0 regardless of the visual target
261 location.

Visual bias as a descriptor of error trend

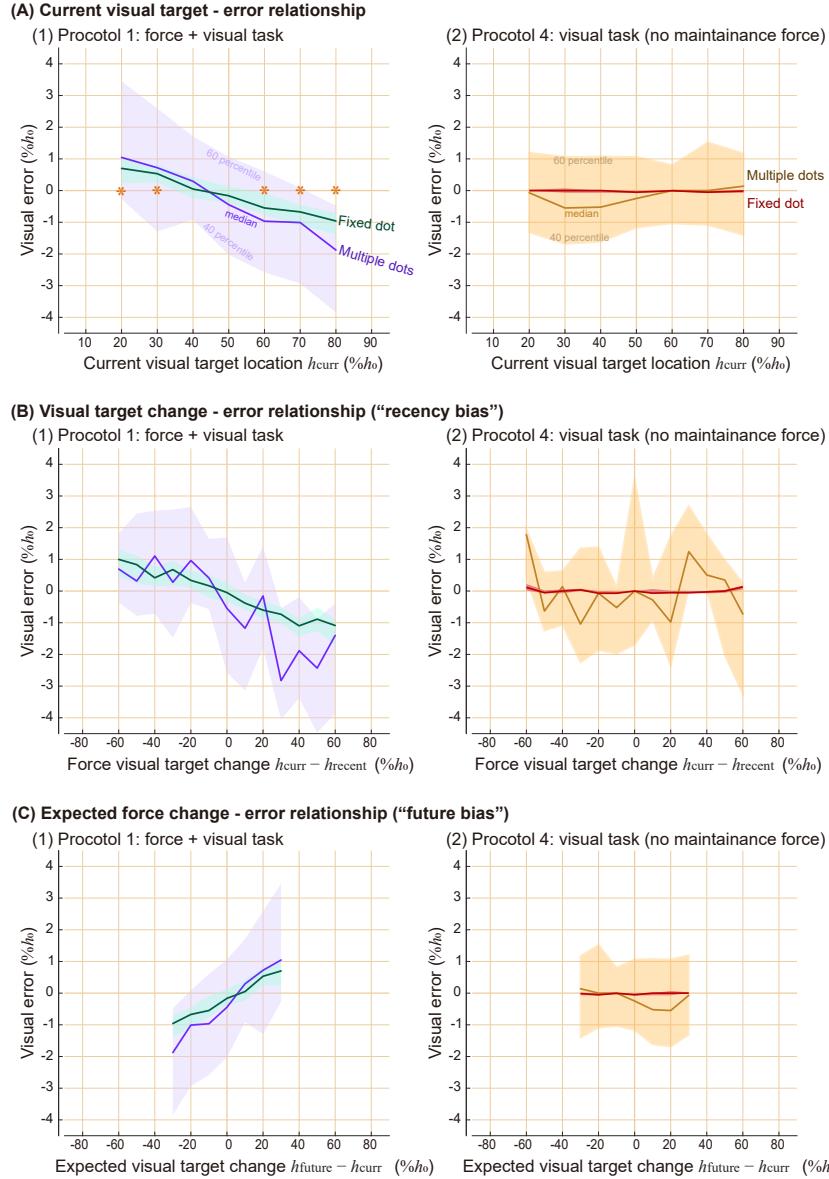


Figure 11. Tracking errors in visual domain, comparing Protocol 1 and Protocol 4. Results from target type of single dot and multiple dots are shown here. Shaded areas represent 40 and 60 percentile data, thicker lines in the middle of the area are median of the data. We represent visual error with respect to (A) current visual target location, (B) distance to the center of the visual target, (C) distance to the recent past target location. Asterisks in top row sub-panel indicate statistically significant difference from zero for both target types separately. There were no significant difference from zero from Protocol 4.

262 Discussion

263 We measured force tracking errors while subjects performed tracking tasks of various protocols
264 and modeled their behaviors using an objective function with an error term and force bias term.
265 Our original aim was to measure the trade-off between task performance and effort, but we needed
266 to include bias term(s) rather than an effort term in the objective function to explain the result. The
267 main discrepancy between the observation and our original expectation was that people made
268 consistent positive force errors in some ranges, which is not explainable by either effort-saving
269 tendency or error-reducing tendency. Central tendency bias and recency bias, which means ten-
270 dency to shift towards the center of the tasks and towards the recent action, seem to explain the
271 data well when combined with an error model. To our knowledge, these biases in force tracking
272 tasks are not previously observed and modeled.

273 In certain modeling scenarios, refining the control objective function, as demonstrated in this
274 study, can yield significant benefits, while in others, the advantages may be less evident. The useful-
275 ness of such refinement may depend on the level of detail at which one examines the phenomenon,
276 and on practical considerations such as computational requirements and analytical solvability. On
277 the other hand, understanding human tendency to make biases while performing motor tasks
278 could have broader applications in various fields, including human-machine interaction
279 and human factors engineering. Additionally, exploring the connection between these biases and
280 neuromuscular disorders could also provide deeper insights into motor control and hold the po-
281 tential for various clinical applications, including diagnosis and rehabilitation.

282 We initially attempted to model the force tracking error using an effort model because it is a
283 common expectation that motor control is usually done in an energetically optimal way. Observa-
284 tion that people apply more force than needed in low-medium force range seem to be an example
285 that people did not optimize for energy, at least during the experiment. However, could it turn
286 out to be energetically cheaper to make such positive biases? For example, if there is a huge cost
287 associated with force changes, it could be more efficient to spend slightly more energy on maintain-
288 ing force and save a bigger energy by reduced a force change. Since we did not directly measure
289 the energetic cost during the experiment, we cannot conclusively claim that people's behavior was
290 energetically suboptimal. However, our speculation is that energy saving is not a primary explana-
291 tion of the biases, because: 1) transient cost people save seem to be small compared to what they
292 spend throughout the sub-trials while maintaining the force, 2) usually "fast change" is regarded
293 energetically costly, but people did not seem to slow down the transient even when there was no
294 explicit speed requirement, suggesting that transient cost is not of a significant consideration, and
295 3) people often did reach a lower force than the target, and increased the force to again to make
296 positive force error at the end, which is shown as followings (**Box 1**).

323 We modeled force tracking error as an outcome of competing goals of reducing error and bias-
324 ing towards a center, but there could be alternative models to capture the same phenomenon. We
325 used our bias models to predict errors from additional protocols, but the model is still descriptive
326 in a sense that we did not test for the underlying mechanism or causality of such bias. Our descrip-
327 tion of biases suggests that people tend to reduce force change, either from the past by making an
328 incomplete shift, or toward the future by making an anticipatory error towards the expected target.
329 However, since we have not proven whether reduction in force change is the primary cause of the
330 errors, it may turn out to be a secondary change following another one. It would be an interesting
331 future work to investigate the causality of the biases.

332 Including more terms in the objective function model could capture a broader range of phenom-
333 ena, but it could also make the inverse optimization problem more underdetermined, potentially
334 leading to an unreliable solution. Central tendency bias and recency bias exhibit similar effects on
335 the overall trend predictions in our experimental designs; thus, one may not need to model both
336 biases to explain some behaviors, depending on their goals. We did not particularly design exper-
337 iments to distinguish these two biases, and there were multiple sets of solutions yielding similarly

298

Box 1. Undershoot in descending sub-trials.

299

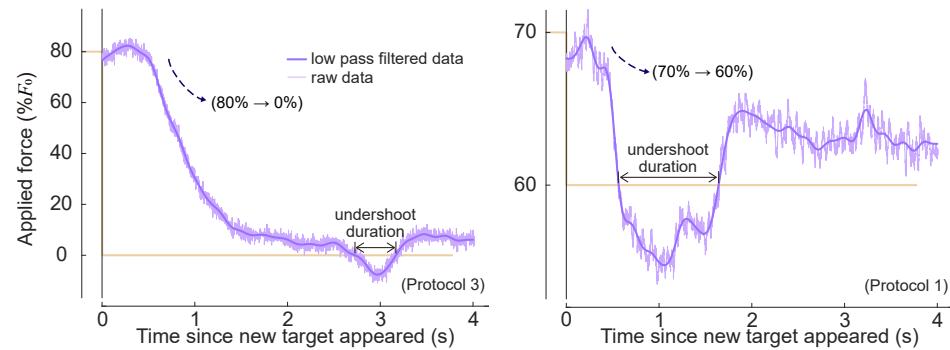
300

301

Undershoot in descending sub-trials (where current target force was lower than the recent past target) are particularly interesting, because it provides insights into the underlying reason of the positive force errors.

302

Descending force transient examples



303

304

305

306

307

Box 1—figure 1. Example force time series during sub-trials that had descending target forces. Dim horizontal lines indicates recent past target and the current target. Lighter purple lines show raw data, and darker purple line show low pass-filtered data that was used in undershoot analysis. We defined undershoot duration as the duration where applied force was lower than the target force during descending sub-trials.

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

Undershoots in descending sub-trials (where current target force was lower than the recent past target) are particularly interesting, because they provide insights into the underlying reason of the positive force errors. Overshoots and undershoots that go past the target were quite common when we look at individual sub-trials, although they were less obvious in averaged force transient responses (figure 9 and figure 10) due to the large variabilities in both amplitude and timing of the responses between sub-trials. In protocols 1, 2, 3 full range and reduced range trials, 72-90% of descending force transients had undershoot of more than 0.2 seconds. Among all descending transients, 67-85% of sub-trials had positive tracking errors at the end. Among descending transients that ended up having positive errors, 68-86% had undershoot of more than 0.2 seconds. There were many sub-trials where the force recovered back to a higher force after an undershoot during descending force transients like the sub-trials shown here, where it did not seem like an unavoidable by-product of a fast responses due to physiological constraints, nor a strategy to save transient energy. This suggests that positive biases on low-medium range target forces are not for energetic advantages, but for some other benefits.

338 good inverse optimization results (figure 7). Our subgroup analysis suggests that both seem to
339 have an effect, but we do not claim that we quantitatively verified recency bias and central ten-
340 dency bias. It would be interesting to design experiments to further disambiguate the biases that
341 originate from the past action and originate from the anticipatory action for the future. For exam-
342 ple, one could design a trial order that cycles through three stages, two random non-zero target
343 forces followed by one zero target, so that average past action is distinct from the average future ac-
344 tions. Even the formulation of the error-reducing tendency must have been an over-simplification
345 if the aim was to closely understand what people perceive as a target. Error function could have a
346 different form depending on the number of dots presented as a target, or could take into account
347 how individual subjects subjectively defined the tasks.

348 There could be various biases in human motor control, and some of them could happen imme-
349 diately without task-specific experiences while some of them may need learning of a task. Some of
350 the biases could take a long time to change its properties, while others may rapidly adapt to each
351 of the tasks. We assumed that subjects readily had an expectation of the center of the target when
352 we modeled central tendency bias, but further research is needed to justify this in depth. We chose
353 to model the central tendency bias this way because we had a speculation that the bias does not
354 seem to rely on learning from each trial, because force tracking error seemed to happen instantly
355 when testing began, and because we do not observe clearly changes over the course of the trial.
356 However, we do not have strong evidence to show whether subjects make biases based on their
357 experience of the trial or expectation of it, or based on the combination of them or even neither
358 of them. It is conceivable that people may start making biases based on their prior expectation
359 and refine the expectation as they have more experience with the task. It would be interesting
360 to design experiments that give subjects a more different impression of the task from its actual
361 composition, and to see which one is a better predictor of the behaviors.

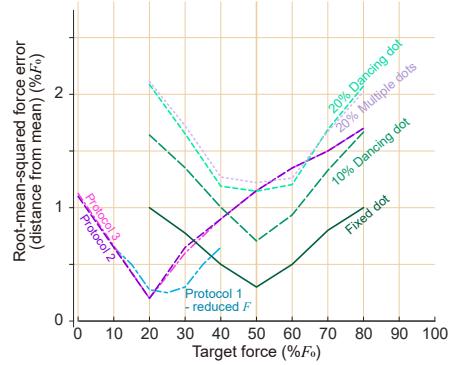
362 One of the reasons that biases in force tracking tasks were not commonly recognized in biome-
363 chanics studies might be due to the specifics of the common experimental designs. In biome-
364 chanics literature, it is common to 1) measure force transients that start from rest, and 2) report
365 absolute value of the error, which omits the directionality of the error. Researchers sometimes
366 describe that people are less accurate at producing bigger forces. Our findings suggests that this
367 description might be highly context-dependent, and there could be a more general description of
368 the phenomenon. Our error-and-bias model predicts that absolute error is minimized near the
369 center of the tasks (figure 12A), rather than monotonically increasing with the force magnitude.
370 Despite that overall root-mean-squared (RMS) errors in experimental results were shifted up com-
371 pared to the model predictions (potentially due to human variabilities), our model captures the
372 overall trend of error (figure 12B) that had minima near the center of the task.

373 Our results are still consistent with the previous studies, when we look at the non-zero sub-trials
374 of our Protocol 2 and 3, which were always followed by a zero force sub-trials. RMS errors indeed
375 roughly increased with the force magnitude in these cases. Directionality of the error during these
376 sub-trials was also consistent with previous studies: we predict and measure negative errors, and
377 previous studies also usually measured negative force tracking errors (e.g., directly reported in
378 *Todorov (2002)*, could be inferred from figures in (*Kudzia et al. (2022)*)). However, what had been
379 usually not measured and reported in previous studies is that RMS force error during zero force
380 sub-trials increases again (figure 12A), and that people tend to make a positive error for this range.
381 The RMS error trend was even more dramatically different from the monotonic increase when we
382 consider Protocol 1 results, where target forces were not reset to zero force.

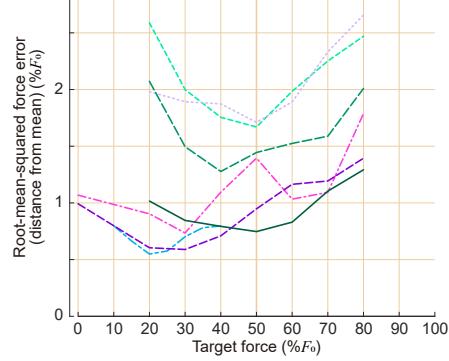
383 This leaves us a question whether we could truly eliminate order effects in scientific measure-
384 ments. As we pointed out, even separating trials by resting periods and having a “fresh start” could
385 give us a wrong impression of the phenomenon, because starting from zero is one of the special
386 cases as well. It would be important to be 1) consistent and clear with what we measure, 2) rec-
387 ognize that what we observe might be a special case of a more general phenomenon, and 3) try
388 to distinguish co-varying variables if possible. For example, force magnitude and force increment

Other measures of force tracking accuracy

(A) Model: root mean squared error



(B) Experimental results: root mean squared error



(C) Experimental results: signal dependent noise

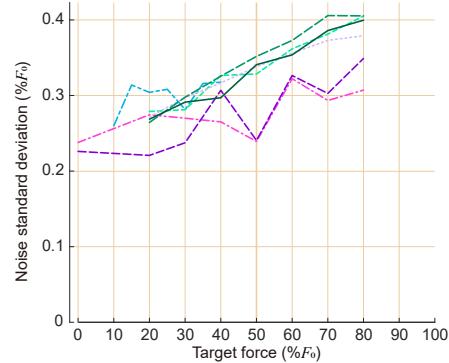


Figure 12. Root-mean-squared error and signal dependent noise during force tracking tasks. For root-mean-squared (RMS) error, we use error defined as a distance between the mean of the target distributions and average force applied during the last 0.5 seconds of the sub-trials. We show RMS error that (A) model predicted and (B) measured from experiments. Lines of different colors indicate different protocols and target types. Target force on the x axis indicates mode of the target distributions. (C) Standard deviation of higher frequency component of the force, with respect to the target force. This higher frequency component of force is referred as motor noise in some literature, and its dependency to the target force is an example of signal dependent noise. Lines of different colors indicate different protocols and target types.

389 co-varies if tests always begin from zero, but they did not co-vary in our Protocol 1, and provided
390 us with a more general view of the nature of force tracking errors. On the other hand, signal de-
391 pending noise, which is defined as standard deviation of force after removing a low frequency
392 component, had a closer match between trials and seemed to increase with force (figure 12C; con-
393 sistent with *Jones et al. (2002)*). This quantity seems to be less context-dependent than force error,
394 and thus, has a higher chance of being explained by simpler physiological properties such as mus-
395 cle characteristics or nervous system dynamics. However, what we measured could still be a partial
396 description of a more general phenomenon, as it always is in science.

397 It was interesting to us that we did not observe a significant bias when people performed an
398 almost fully visual task (figure 11), considering that many studies on recency bias and central ten-
399 dency bias are measured from visual pattern recognition tasks. It is conceivable that human minds
400 would have various biases without involving the muscular system. On the other hand, some biases
401 may be caused by more passive and physical properties of the body without involving perceptions
402 and judgements, and some biases might be caused by combinations of them. From studying the
403 biases of force tracking tasks, we became curious if there is a chance that some of the percep-
404 tual biases could be attributed to motor control biases more than researchers commonly thought.
405 For example, ocular motor control may have a similar bias as what we observed in this study that
406 involved arm muscles, and that could result in a bias in visual sensation. There could be biases
407 caused during measurement procedures as well, because subjects were often asked to indicate
408 their perception of patterns using some physical device by manipulating it, which involves some
409 degree of motor control.

410 Force bias does not seem to be a commonly recognized phenomenon in an isolated motor con-
411 trol setting, but we could relate it to a broader field of motor control. In a singing study, it was
412 reported that people tend to compress the pitch shifts, and "bad singer"s tend to make more such
413 compression (*Pfordresher and Brown, 2007*). In speech studies, there is a window model suggest-
414 ing that people make a minimal shifts between acceptable windows of some variables (*Keating
415 et al., 1990*). Moreover, when we consider the recency bias in broader terms that people tend
416 to keep acting similarly, there are more such examples in biomechanics studies. For example,
417 walk-to-run and run-to-walk speeds are reported to be different ("hysteresis effect in speed for the
418 walk-run transition", *Diedrich and Warren Jr (1995)*), which is not simply explained by the idea that
419 animals choose the most economic gait at a given speed. People change their behavioral mode
420 between one-handed grasping and and two-handed grasping depending on the size of the object,
421 and it is reported that there is an order effect in such shifts (*Frank et al., 2009*). Some researchers
422 attribute these biases to "economy" that people save energy by reducing transitions, although this
423 claim is usually difficult to be directly tested. We hope that an extension of our current study could
424 give insights into these broader fields of behavioral studies.

425 To conclude, we examined which objective function best describes human force tracking errors.
426 In many motor control studies, objective functions are modeled to include error and effort terms,
427 typically in quadratic forms. However, we designed an experiment to test them and indeed found
428 a different formulation. Exponent on the error we found was smaller than 2, and bias term(s) were
429 needed instead of an effort term in order to capture the behavior that was consistent across par-
430 ticipants. Our findings on biases suggest that biological motor control, even in a simple isometric
431 force production task, can be highly context-dependent, and commonly hypothesized formulation
432 might not be adequate to represent the phenomenon. As we continue to explore the intricacies
433 of human force control, further research may offer valuable insights into our understanding of
434 human movement behaviors and underlying principles of the nervous system.

435 Methods

436 Experimental procedure

437 Subjects ($N_{\text{subjects}} = 12$; 4 Female, 8 Male) participated in Protocol 1, and a subset of them participated in additional studies later. This human subject research was reviewed and approved by
438 The Ohio State University's Institutional Review Board, and all subjects participated with written
439 informed consent.

440 We measured elbow height from the ground and subtracted half of the subject's forearm length
441 to set a platform height they pressed onto, which made the forearm angle about 30 degrees below
442 the horizontal line. After setting the force platform, we measured the individualized force target
443 range F_0 by asking the participants to apply a force that they can comfortably hold for 30 seconds.
444 We encouraged participants to apply reasonably high force, and reduce it until it felt comfortable
445 to hold. Real time force feedback was given on the monitor, and we encouraged them to control
446 the feedback bar location. After measuring the force range, we verified it again that subjects could
447 comfortably exert the force for 30 seconds. The process was repeated until we found the force
448 that the subject was comfortable with producing for an extended amount of time.

449 After giving participants verbal description of the tasks, we provided them with practice trials,
450 which were shorter versions of the trials so that they can see the test procedure and try the tasks.
451 We introduced different target types in turn – fixed single dot, then single dancing dot, and then
452 multiple dots. Target changed color between blue and red each sub-trials, and all participants
453 confirmed that the colors were significantly different for them. We instructed the participants that
454 the target location changed every 4 seconds with the change of the target color, and asked them to
455 apply force so that the force feedback bar points the target. For vague targets, we told them that
456 the dots are dancing around some fixed point and are not moving away as long as the color stays
457 the same, and asked them to point where they perceive as a target from the overall impression of
458 the dots. We explicitly asked them not to chase individual occurrence of the dots, but to perceive
459 its center and place the bar there. Participants practiced each of the target types for at least 5
460 sub-trials, and we repeated the practice if subjects desired to do so or were unclear with the task.

461 Each trial was composed of 74 sub-trials. The first three trials were excluded from the analysis to
462 make sure that subjects adapt to a different target type. Following 70 conditions covered possible
463 combinations of force level and distribution parameters in a random order, which is described in
464 the following paragraph. After these 70 sub-trials that are used for analysis, the last sub-trial of
465 each trial was 100% F_0 one, from which we checked that subjects were still capable of producing
466 more than 80% of F_0 . 4 seconds of 74 sub-trials took about 5 minutes.

467 We tested 20, 30, ..., 80% of F_0 in Protocol 1, and used distribution parameter $\alpha = -2, -1, -0.5,$
468 $-0.1, -0, +0, +0.1, +0.5, +1, +2$ to generate vague targets, resulting in 7 force levels \times 10 distribution
469 parameters = 70 combinations. For the fixed single dot condition, we measured each force level
470 10 times to ensure the same number of sub-trials. In Protocol 1 - reduced F , force levels we tested
471 were 10, 15, ..., 40% F_0 . In Protocol 2 and 3, non-zero forces were 20, 30, ..., 80% F_0 , and each
472 force level was measured 5 times, again resulting in a total of 70 sub-trials including zero force
473 periods. In Protocol 2 and 3 - reduced F conditions, non-zero force levels were 20, 25, 30, ..., 50
474 % F_0 . Protocol 3 had a binary visual target locations, which were same as the target for 0 force and
475 80% F_0 in Protocol 2.

476 We defined distribution parameter α as follows. Positive non-zero distribution parameters cor-
477 responds to a shape parameter of Pareto Distribution truncated between 0 and 5 (**Zaninetti and**
478 **Ferraro, 2008**). The range was scaled to match 10% and 20% F_0 . We generated distributions of neg-
479 ative non-zero α same as a positive α , except that the whole distribution was flipped so that the tail
480 of the distribution goes the opposite way. Two symmetric distributions were also used: truncated
481 uniform distribution was referred to as $\alpha = -0$, and truncated normal distribution that had similar
482 standard deviation as other Pareto distributions were referred as $\alpha = +0$ in this study as special
483 cases of skewed distributions.

485 Six of the subjects had a force feedback that went down when they applied force, and six of
486 the subjects had a force feedback that went up when they applied force. We did so to see if there
487 is a visual bias linked to a vertical directionality. We did not observe a notable difference between
488 subjects of each feedback directionality. Each time before the test started, we asked subjects to
489 place their hands on the platform and relax arm muscles, try not to exert active force onto the
490 platform and let the platform take the weight of the hands. This hand weight was set to zero force,
491 so that participants did not need to actively spend energy lifting up their hands to match low force
492 requirements.

493 Protocol 1 and 4 involved testing vague targets. In Protocol 1 – dancing dot target condition,
494 a dot appeared within 10% and 20% F_0 range, and in Protocol 1 and 4 – multiple dots conditions,
495 240 dots appeared within 20% F_0 . We tested all target types in a random order two times. Other
496 protocols were measured once per protocol. Subjects took about 2 minutes of rest between trials
497 and occasionally took longer breaks as desired.

498 **Data collection and processing**

499 Custom MATLAB software was used to show targets and force feedback on the screen. Force plate
500 (Bertec Corporation, Ohio, USA) mounted on the ground was used to collect force data. Force
501 plate data was collected through a motion capture interface (Nexus, Vicon, Oxford, UK) that was
502 relayed to MATLAB software. Target and force feedback were updated at 60 Hz, which matched
503 the monitor refresh rate. Force was collected at 8 times faster speed, which is 960 Hz. Multiple
504 dots changed their horizontal positions every frame. The target radius was about 0.7% F_0 .

505 Mean value of force during the last 0.5 seconds of each sub-trials was defined as the steady-
506 state force value, and this value minus the target force was defined as force error. We normalized
507 force error by F_0 . Mean value of force during the first 0.1 seconds of each sub-trials was defined
508 as an initial force, and was used to calculate recency bias. To study signal dependent noise, we
509 obtained a fast component of the force change by subtracting a slow frequency response from
510 a faster frequency response. Fifth order Butterworth low pass filters with cut-off frequency of
511 25Hz and 5Hz were applied to force data using zero phase filtering in order to calculate these slow
512 and faster responses. Standard deviation of this fast component during the last 0.5 seconds was
513 reported as a "noise" to investigate signal dependent noise (figure 12).

514 **Protocol 4 - test using computer mouse with negligible forces**

515 Custom MATLAB software was used to conduct tests of Protocol 4. We kept the testing interface as
516 similar as possible to other protocols, except that vertical position of the horizontal bar that used
517 to indicate the applied force was changed to indicate the vertical position of the mouse pointer on
518 the screen. The cursor was hidden and subjects could only see the height of the horizontal bar as
519 they moved the computer mouse.

520 We used a MATLAB callback function that responded to mouse position change. Trials were
521 similar to those of Protocol 1, in terms of test duration and conditions. We tested a fixed single
522 dot and multiple dots using this interface. Unlike Protocol 1, multiple dots did not change their
523 horizontal locations in Protocol 4, because we could not guarantee a constant refresh rate through
524 the interface we employed. We recorded the time and position of the mouse pointer each time
525 there was a change. Clicking was not required to perform the task.

526 **Predictions for additional protocols based on the bias terms.**

527 We considered two types of biases in this paper. Predictions of the force tracking error based on
528 central tendency bias and recency bias are as follows.

529 **Predictions based on central tendency bias.** We modeled a tendency to shift towards the "cen-
530 ter" of the tasks, where "center" was defined as the mean of the target forces throughout the
531 trial. The mean of the tasks in Protocol 1 was 50% F_0 , and the mean of the tasks in Protocol

532 1-reduced F was 25% F_0 . Using this model, we expected that errors would in general cross 0
533 around these force levels in each of the protocols. In Protocol 2 and 3, since half of the sub-
534 trials were at 0%, the mean of the targets was 25% F_0 . We expected that zero force sub-trials
535 would have positive errors, and non-zero force sub-trials would have more negative errors
536 as the target force increased. We used the same center for Protocol 2 and 3 – reduced F ,
537 because subjects were unaware of the reduced force range.

538 **Predictions based on recency bias.** We modeled a tendency to shift towards the recent past ac-
539 tion. In Protocol 1 and Protocol 1-reduced F , prediction based on recency bias model was
540 that lower target forces would have positive errors, because their recent past force is higher
541 than their current forces on average. Similarly, the prediction on higher forces was that they
542 would have negative errors, and medium range forces would have errors near zero. This dis-
543 tinction of low, medium, high forces were all relative to the force range of each trials in this
544 bias model. Similar to the central tendency bias model, we expect that errors would cross
545 zero near 50% F_0 and 25% F_0 respectively for Protocol 1 - full range and reduced F conditions.
546 In Protocol 2 and 3 where there were zero reset periods, we expected that zero force sub-
547 trials would have positive errors because their recent past forces were always higher than
548 the current force. Similarly, we expected that non-zero sub-trials would have more negative
549 error as target force increased, because their recent past forces were always zero. Protocol
550 2 and 3 – reduced F were expected to have similar error trends as their full F range versions.
551 Only difference could be on zero force sub-trials, because their average distance to the recent
552 past target was smaller in reduced range protocols than in full-range protocols, thus could
553 result in a slightly smaller positive error on zero force sub-trials.

554 Both central tendency bias and recency bias models predict similar overall error trends because
555 their expected values are similar to each other. Both models predict that neither error-target force
556 relationship, nor visual error-visual location relationship will be consistent across different proto-
557 cols, because bias is dependent on force and the context of the experiments. In addition, both bias
558 models predict that there is a bias in the force domain, not in the visual domain; thus, we predicted
559 that Protocol 4 will not have such bias because the task does not require maintenance force.

560 **Inverse optimization**

561 We did grid search to perform inverse optimization to select hyperparameters of objective function
562 models. We used a similar mathematical formulation to represent both central tendency bias and
563 recency bias had, which was:

$$F^* = \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^p}{N} + bd_{\text{bias}}^\beta,$$

564 where F^* is the prediction of the force which minimizes the given objective function, error_j is
565 absolute distance between j -th dot and the force applied, summed over number of target dots
566 N . d_{bias} is absolute distance from the applied force to the bias center. It is the distance to the
567 perceived average of the target forces for central tendency bias, and distance to the initial force at
568 the beginning of the sub-trial for recency bias.

569 We searched for error exponents p between 1.4 and 1.9 range by the increments of 0.1, bias
570 weightings b between 0.01 and 0.25 by the increments of 0.01, and bias exponent β between 1
571 and 2 by the increments of 0.05. For each combination of these hyperparameters, we calculated
572 the model prediction of each sub-trials. The model predicts the forces that minimize the given
573 objective function. To avoid local minima issue and to improve the computational efficiency, we
574 evaluated the objective function within $\pm 30\% F_0$ around the target by the increments of 0.05% F_0 ,
575 and found the minimum value among them.

576 After calculating predictions for each sub-trials and each combinations of hyperparameters, we
577 selected the set of hyperparameters that produced similar error-force relationship as the experi-
578 mental results. Since error trend we aimed to model was in a relatively small magnitude compared

579 to large inter-subject and intra-subject variabilities, we selected an objective function that captures
580 the overall behavior. The objective functions that minimized mean RMS error between mean data
581 and mean model prediction on each force condition for the entire protocols were:

582 • Error-and-central tendency bias model:

$$F^* = \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.7}}{N} + 0.09d_{\text{ctr-bias}}^{1.55},$$

583 • Error-and-recency bias model:

$$F^* = \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.5}}{N} + 0.15d_{\text{rec-bias}}^{1.3}.$$

584 After performing inverse optimization on central tendency bias model and recency bias model
585 separately, we confirmed that they indeed have a similar formulation and have a similar effect
586 of the force error predictions. We performed an inverse optimization on a combined bias model
587 while keeping some of the hyperparameters fixed based on these results. We chose to fix some
588 parameters to avoid overfitting, because we have confirmed that two biases have similar effects
589 on the overall results, and human data is already very noisy.

590 We scanned the error exponent p between 1.4 and 1.9 by the increments of 0.1, while changing
591 two bias weightings between 0.01 and 0.25 by the increments of 0.01. The bias exponent β
592 was fixed to the optimal value that was found earlier. The objective function that minimized RMS
593 distance between mean data and mean model predictions on error-force relationships was:

594 • Error-and-combined bias model:

$$F^* = \arg \min_F \sum_{j=1}^N \frac{\text{error}_j^{1.6}}{N} + 0.05d_{\text{ctr-bias}}^{1.55} + 0.07d_{\text{rec-bias}}^{1.3}.$$

595 References

596 **Baron S**, Kleinman DL. The human as an optimal controller and information processor. IEEE Transactions on
597 Man-Machine Systems. 1969; 10(1):9–17.

598 **Berret B**, Chiovetto E, Nori F, Pozzo T. Evidence for composite cost functions in arm movement planning: an
599 inverse optimal control approach. PLoS computational biology. 2011; 7(10):e1002183.

600 **Dao TK**, Pan TS, Pan JS. A multi-objective optimal mobile robot path planning based on whale optimization
601 algorithm. In: 2016 IEEE 13th international conference on signal processing (ICSP) IEEE; 2016. p. 337–342.

602 **De Groot F**, Falisse A. Perspective on musculoskeletal modelling and predictive simulations of human move-
603 ment to assess the neuromechanics of gait. Proceedings of the Royal Society B. 2021; 288(1946):20202432.

604 **De Groot F**, Kinney AL, Rao AV, Fregly BJ. Evaluation of direct collocation optimal control problem formulations
605 for solving the muscle redundancy problem. Annals of biomedical engineering. 2016; 44:2922–2936.

606 **Diedrich FJ**, Warren Jr WH. Why change gaits? Dynamics of the walk-run transition. Journal of Experimental
607 Psychology: Human Perception and Performance. 1995; 21(1):183.

608 **Emken JL**, Benitez R, Sideris A, Bobrow JE, Reinkensmeyer DJ. Motor adaptation as a greedy optimization of
609 error and effort. Journal of neurophysiology. 2007; 97(6):3997–4006.

610 **Frank TD**, Richardson MJ, Lopresti-Goodman SM, Turvey MT. Order parameter dynamics of body-scaled hys-
611 teresis and mode transitions in grasping behavior. Journal of biological physics. 2009; 35:127–147.

612 **Izawa J**, Rane T, Donchin O, Shadmehr R. Motor adaptation as a process of reoptimization. Journal of Neuro-
613 science. 2008; 28(11):2883–2891.

614 **Jin R**, Rocco P, Geng Y. Cartesian trajectory planning of space robots using a multi-objective optimization.
615 Aerospace Science and Technology. 2021; 108:106360.

616 **Jones KE**, Hamilton AFdC, Wolpert DM. Sources of signal-dependent noise during isometric force production.
617 *Journal of neurophysiology*. 2002; 88(3):1533–1544.

618 **Kalakrishnan M**, Pastor P, Righetti L, Schaal S. Learning objective functions for manipulation. In: *2013 IEEE*
619 *International Conference on Robotics and Automation* IEEE; 2013. p. 1331–1336.

620 **Kalman RE**, et al. Contributions to the theory of optimal control. *Bol soc mat mexicana*. 1960; 5(2):102–119.

621 **Keating PA**, Kingston J, Beckman ME. The window model of coarticulation: articulatory evidence. *UCLA Working*
622 *papers in Phonetics*. 1990; 69:3–29.

623 **Kleinman DL**, Baron S, Levison W. An optimal control model of human response part I: Theory and validation.
624 *Automatica*. 1970; 6(3):357–369.

625 **Körding KP**, Wolpert DM. The loss function of sensorimotor learning. *Proceedings of the National Academy of*
626 *Sciences*. 2004; 101(26):9839–9842.

627 **Kudzia P**, Robinovich SN, Donelan JM. Characterizing the performance of human leg external force control.
628 *Scientific Reports*. 2022; 12(1):4935.

629 **Lee LF**, Umberger BR. Generating optimal control simulations of musculoskeletal movement using OpenSim
630 and MATLAB. *PeerJ*. 2016; 4:e1638.

631 **Liu S**, Sun D. Minimizing energy consumption of wheeled mobile robots via optimal motion planning.
632 *IEEE/ASME Transactions on Mechatronics*. 2013; 19(2):401–411.

633 **Mi Z**, Yang Jj, Abdel-Malek K. Optimization-based posture prediction for human upper body. *Robotica*. 2009;
634 27(4):607–620.

635 **Miao C**, Chen G, Yan C, Wu Y. Path planning optimization of indoor mobile robot based on adaptive ant colony
636 algorithm. *Computers & Industrial Engineering*. 2021; 156:107230.

637 **Pellegrinelli S**, Borgia S, Pedrocchi N, Villagrossi E, Bianchi G, Tosatti LM. Minimization of the energy consump-
638 tion in motion planning for single-robot tasks. *Procedia Cirp*. 2015; 29:354–359.

639 **Pfordresher PQ**, Brown S. Poor-pitch singing in the absence of "tone deafness". *Music Perception*. 2007;
640 25(2):95–115.

641 **Srinivasan M**. Fifteen observations on the structure of energy-minimizing gaits in many simple biped models.
642 *Journal of The Royal Society Interface*. 2011; 8(54):74–98.

643 **Todorov E**. Cosine tuning minimizes motor errors. *Neural computation*. 2002; 14(6):1233–1260.

644 **Zaninetti L**, Ferraro M. On the truncated Pareto distribution with applications. *Open Physics*. 2008; 6(1):1–6.

645 **Acknowledgments**

646 HXR was supported by the U. Calgary Department of Biomedical Engineering. MS was supported
647 in part by NIH-R01GM135923-01 and NSF SCH grant 2014506.